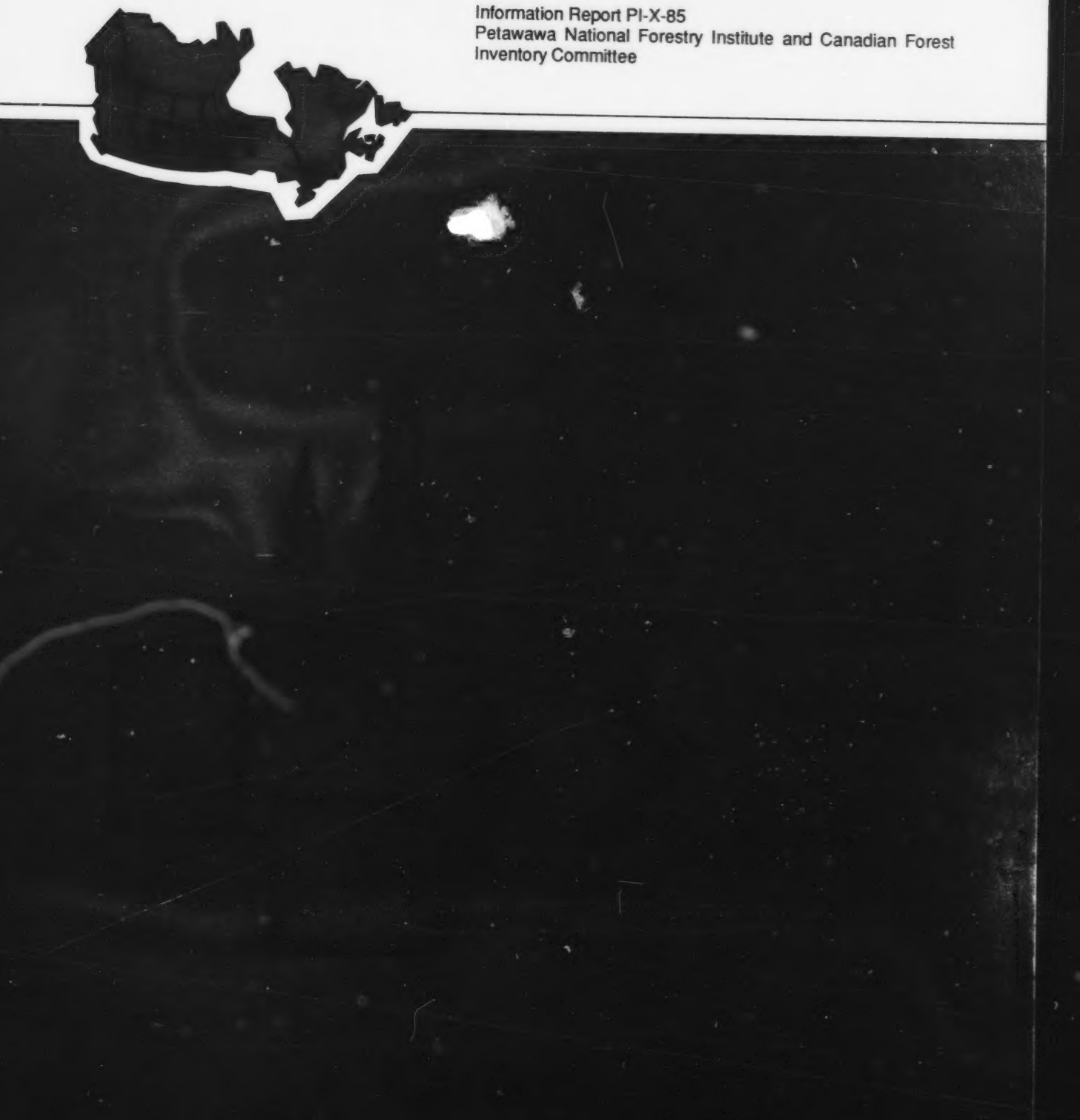
Government of Canada | Gouvernement du Canada
Canadian Forestry Service | Service canadien des forêts

Theory and methodology of single tree volume estimation

S.J. Titus and D.J. Morgan

Information Report PI-X-85
Petawawa National Forestry Institute and Canadian Forest Inventory Committee



89-010307

PETAWAWA NATIONAL FORESTRY INSTITUTE

In common with the rest of the Canadian Forestry Service, the Petawawa National Forestry Institute has as its objective the promotion of better management and wiser use of Canada's forest resource to the economic and social benefit of all Canadians. Objectives of program activities carried out at the Institute support this goal through discovery, development, demonstration, implementation, and transfer of innovations. Because it is a national institute, particular emphasis is placed on problems that transcend regional boundaries or that require special expertise and equipment that cannot be duplicated in CFS regional establishments. Such research is often performed in close cooperation with staff of the regional centres, provincial forest services, and the forest industry.

Research initiatives and technical services at the Institute encompass five major activities:

FOREST GENETICS AND BIOTECHNOLOGY — Integrates projects in tree genetics, soil microbiology, micropropagation, molecular genetics, meteorology, and seed research. It also includes the client services and seed bank operations of the National Tree Seed Centre, a long-standing program with extensive international affiliations.

FOREST MANAGEMENT SYSTEMS — This program integrates projects in fire, remote sensing, modelling, growth and yield, and forest pest management to provide research and development for the formulation and demonstration of forest management systems.

NATIONAL FOREST RESOURCE STATISTICS — Provides biological, technical, and socioeconomic information on Canada's forest-based resources. The program involves progressive development of databases, such as FIDSINFOBASE, and establishment of new databases and software in support of policy development in forestry. The Forest Inventory Program collates information on the forest resource at a national level, maintains the Canadian Forest Resources Data System, and prepares the national forest inventory.

COMMUNICATIONS — Integrates activities of the library, public awareness, information, and editing and publications projects. The Institute is visited by more than 20 000 people every year. There is a Visitor Centre for the public, self-guided tours, and an extensive education project. The national repository of all scientific and technical publications of the CFS and the principal CFS publications distribution centre are both located at PNFI.

THE RESEARCH FOREST — Besides natural stands manipulated in a variety of ways for silvicultural research, the 98 km^2 Petawawa Forest contains extensive areas of plantations dating back six decades. Research plantations are a source of growth and yield data derived from cultural experiments, and they are becoming valuable for pedigreed genetic materials for micropropagation and molecular genetics studies. The forest also offers opportunities for short- and long-term testing of forest management strategies.

THEORY AND METHODOLOGY OF SINGLE TREE VOLUME ESTIMATION

S.J. Titus, Department of Forest Science, University of Alberta
D.J. Morgan, Timber Management Branch, Alberta Forest Service
Alberta Forestry, Lands, and Wildlife

Information Report PI-X-85
Joint Publication No. 1
Petawawa National Forestry Institute
and Canadian Forest Inventory Committee
1988

©Minister of Supply and Services Canada 1988
Catalogue No. Fo46-11/85-1988
ISBN 0-662-56318-2
ISSN 0706-1854
Printed in Canada

Copies of this publication may be obtained free of charge from the following address:

Publications Distribution Centre
Petawawa National Forestry Institute
Chalk River, Ontario
K0J 1J0

Telephone: 613-589-2880

A microfiche edition of this publication may be purchased from:

Micromedia Ltd.
Place du Portage
165, Hôtel-de-Ville
Hull, Québec
J8X 3X2

Canadian Forest Inventory Committee (CFIC)

The CFIC is a national committee constituted to address the continuing state, practice, and research in Canada of forest inventory, inventory update, and other forest management statistics. Its mandate includes operational procedures, measurement standards, terminology, and data acquisition, handling and reporting.

Voting membership is comprised of the heads of forest inventory in the provincial and territorial forestry agencies and the Canadian Forestry Service, along with an independent charter member. Other forest inventory expertise is represented through nonvoting membership.

Contents

		Page
Abstract		iv
1.	Introduction	1
2.	Theory of tree volume relations	1
	2.1 Variables	1
	2.2 Tree volume relations	1
3.	Methodology for evaluating tree volume relations	2
	3.1 Population specifications	2
	3.2 Sampling methods	2
	3.3 Measurement procedures	3
	3.4 Curve fitting	4
4.	Using tree volume relations	6
	4.1 Measurement specifications and their impact on usage	6
	4.2 Extrapolation	6
	4.3 Adjustment of existing relationships	7
5.	Conclusion	7
6.	References	7

Abstract

A review of theory and methodology of single tree volume estimation is presented. Guidelines are discussed for dealing with (1) population definition, (2) sampling procedures, (3) measurement methods, and (4) curve fitting techniques.

1. Introduction

Volume is a primary variable used to describe the size and product potential of individual trees and is the basis for developing volume estimates for stands. The purpose of this paper is to summarize the current knowledge about major relationships used to predict individual tree bole volumes. Particular emphasis is given to methodology for evaluating these relations, including both measurement and statistical considerations. The aim is to provide a discussion that will aid the practicing forester faced with developing or evaluating tree volume estimation relationships. This paper is an expansion of an earlier draft checklist for measurement procedures developed by the Mensuration Subcommittee of the Canadian Forest Inventory Committee. A number of related, peripheral topics are not included in this discussion, including defect, quality, merchantability, taper relations, and the definition of the bole in delinquescent (usually hardwood) trees.

2. Theory of Tree Volume Relations

Tree volume may be measured using any of several direct measurement procedures or estimated indirectly based on the relationship of volume to other tree characteristics. Most commonly, the direct approach means cutting trees into sections and summing estimates of section volumes to obtain the tree volume. Measurement of standing trees is sometimes possible using instruments such as an optical dendrometer to obtain measurements of diameter and height at different points along the length of the bole.

While section volumes are ideally obtained using liquid displacement, it is seldom considered in practice. More commonly, section volumes are obtained by application of a volume formula such as Smalian's which assumes sections are frusta of paraboloids. Since direct approaches are both costly and time-consuming, volumes for standing trees are often estimated based on the relationship of volume to easily measured tree characteristics such as diameter, height, and form.

2.1 Variables

The dependent variable is gross tree bole volume in cubic metres, inside bark, from ground level to the tip of the tree. A number of independent variables are related to tree volume, including total height (m) and diameter (cm) at 1.3 m height. In addition, other variables such as species and region may be "fixed" by constructing separate equations for each combination to reduce population variability. Occasionally, other variables such as crown ratio and a measure of tree form may be considered. However, where relationships are used to predict tree volume for routine inventory work, it may be necessary to restrict independent variables to those normally measured and recorded as part of the inventory.

2.2 Tree Volume Relations

In considering the estimation of tree volume, it is useful to model the main stem (bole) as a cone. Volume of a cone is obtained as the integration of the cross-sectional area, based on linear taper, from the base to the tip of the cone producing the formula

$$V = (H / 3) \times \pi (D / 200)^2 \qquad [1]$$

where volume (V), height (H), and diameter (D) are expressed in cubic metres, metres, and centimetres, respectively. This simple geometric model of tree volume is the basis for the most successful relationships used to predict standing tree volume. Variations on this model are obtained

by generalizing the underlying taper model to allow for other taper forms including the paraboloid and neiloid.

In the following discussion three examples are given for both standard and local volume relations. Other model forms have been investigated and reported (Clutter et al. 1983; Husch et al. 1982; Spurr 1952). In this discussion the terms "model," "formula," and "equation" are synonymous.

Three standard volume relations have been used to estimate volume, and all are based on variations of the cone volume formula:

$$V = a + b \times D^2 \times H \quad [2]$$
$$V = a \times D^b \times H^c \quad [3]$$
$$\log V = a + b \times \log D + c \log H \quad [4]$$

where a, b, c, and d are parameters of the models and log refers to the logarithm (traditionally base 10). The third model (Schumacher and Hall 1933) is the linearized form of the second model, which in turn is a slight generalization of the first (Spurr 1952).

Local volume relations typically use only dbh to predict volume and as a result are subject to larger errors:

$$V = a + b \times D^2 \quad [5]$$
$$V = a \times D^b \quad [6]$$
$$\log V = a + b \log D \quad [7]$$

The second model is a generalization of the first, allowing the power of diameter to be estimated; and the third is the logarithmic form of the second model. The estimation of volume from photo-measured tree height in remote sensing applications is also sometimes referred to as a local volume relation.

3. Methodology for Evaluating Tree Volume Relations

This discussion is subdivided into sections considering (1) population specifications, (2) sampling methods, (3) measurement procedures, and (4) curve fitting. Even though these are presented sequentially as separate topics, they are interrelated and parts of the discussion necessarily refer to aspects of other topics because of their importance and impact. For example, both sampling methods and measurement procedures are affected by the assumptions and requirements of curve fitting methodology.

3.1 Population Specifications

Careful specification of the population is required so that the limits of inference are known. Often this means limiting the scope of the relationship to a particular species and geographic region. Not only the species and region of interest must be specified, but also the characteristics of individual trees. Should forked and broken-top trees be included? Should the population be limited to trees younger than the typical rotation age for the species? Answers to these questions depend on whether predictions are meant for ideal trees grown in plantations or the existing natural population.

3.2 Sampling Methods

Sampling for evaluating relationships (regression) is often quite different from sampling for inventory. Here it is important to sample over the entire range of independent and dependent variables, and this may mean a less random, more selective or representative sampling procedure.

Ideally, sample trees should be selected from random locations throughout the population area, but to reduce costs, several trees are often measured at each sample location. However, if the number selected at each location becomes too large in relation to the number of locations there may not be sufficient coverage of population conditions since trees growing near each other tend to be similar. West et al. (1984) have shown that multiple observations from the same sampling unit, such as plots, can lead to biased estimates of variances in testing of regression model hypotheses.

A reasonable number of sample trees are needed, but exactly how many depends on population variability, the strength of relationships, the magnitude of acceptable errors, and cost limits (Demaerschalk and Kozak 1974, 1975; Gove et al. 1982; Marshall and Demaerschalk 1986). A minimum of 30 sample trees will give an indication of the basic nature of the relation, but a larger sample of 100 or more may be desirable depending on the variability and range of tree sizes. In addition, allocation of samples over the range of independent and dependent variables should be uniform only if the variance of errors is the same for different magnitudes of the independent variables. If error variance increases with increasing values of the dependent variable, as it usually does with tree volume, then allocation should be in proportion to the variance and the probability of sampling large trees should be larger (as it is in variable radius or "angle-gauge" plot sampling). Data of this type may require a weighted analysis (see later discussion). Deciding on an adequate sample size is a difficult task. It should be kept in mind that more is not automatically better; quality and distribution of the data are also important.

3.3 Measurement Procedures

Measurement of sample trees requires consistent procedures and minimization of errors, since subsequent analysis of data to evaluate relationships is based on the assumption that measurement errors are negligible.

Diameter measurements are commonly made with calipers or a diameter tape. If the former is employed, two measurements should be made at right angles to each other and the geometric mean used as the average diameter.

Height should be based on measurement of felled trees or precise measurement of standing trees to minimize errors. Common inventory procedures such as use of a clinometer and paced or taped distances are usually not sufficiently precise (Titus and Morgan 1985).

If trees are cut and sectioned to determine volumes, the lower cut may be at ground level or stump height (0.3 m), or both, followed by a cut at the breast-height point (1.3 m). Prior to felling, the distance on each tree from ground to breast height and stump height should be measured and marked. A measurement of diameter outside bark at these lower points should be made to ensure that damage during felling does not affect those dimensions. Marking also assures that these critical points are consistently located. Section lengths above 1.3 m height should be less than 2.0 m to minimize errors in section volumes due to variation in taper. Bark thickness in addition to diameter, or both inside bark (ib) and outside bark (ob) diameters, may be measured. Local stem distortions like burls and branch whorls may be avoided by adjusting section lengths as necessary.

In practice, procedures range from taking fixed, short section lengths (for example, of 1.0 m) to selecting a fixed number of sections (for example, 10) for every tree regardless of height. Fixed short

sections provide better estimates of volume, especially where extreme variations in taper are common, but this practice can be costly for tall trees. The latter approach reduces cost, but may yield poor estimates of section volume (with 10 sections per tree, a 30-m tree will yield section lengths of almost 3 m). Selection of an appropriate compromise requires balancing cost with desired accuracy. The following measurement standards are suggested:

section diameters	— 0.1 cm
section lengths	— 0.01 m
volume	— 0.001 m^3

The following form assumptions are recommended for estimating volume of individual sections:

stump	— a cylinder of neiloid frustum
middle	— paraboloid frustum
tip	— a cone

Interpolations within sections, if required, should employ the same form assumptions. It should be noted that as section length is shortened, the difference in diameter (area) at each end decreases and volume differences between formulas become negligible. Conversely, as section length increases volume differences between formulas increase and the log form assumptions become much more important. The recommendation to use Smalian's formula for the middle sections of a tree is based on the additional assumption of short sections since it overestimates volume on long sections (Husch et al. 1982).

All data should be screened before statistical analysis to detect errors or "outliers." Errors should be corrected if possible; outliers should be discarded if it is determined that there is some obvious error which cannot be corrected. Data points should not be deleted simply because they seem unusual. Graphical data summaries are helpful in identifying outliers and may also aid in preliminary evaluations of the nature of relationships among diameter, height, and volume.

3.4 Curve Fitting

Curve fitting, or regression analysis, is normally divided into two types, linear and nonlinear, according to the model form used. Linear models have been favored for a number of reasons, including better developed statistical theory and relative ease of the curve fitting procedures. With greater computational power available, nonlinear models have increased in popularity because of their ability to provide model forms that match theoretical expectations. The nonlinear form may also be viewed as a generalization of the linear model, which allows fitting of both linear and nonlinear models.

Regression models allow for residual errors with the assumption that the long run average of the errors is zero, errors are independent, and the variance of errors is constant at all levels of the independent variables. In linear regression the added assumption is made that the errors are normally distributed.

These assumptions are often ignored with the result that inappropriate data and analysis procedures are used. Several types of problems may occur, including most commonly:

1. poor fit of the model (often expressed as bias),
2. unequal variance of errors (typically increasing as the independent variable increases),
3. correlated errors (multiple measures on the same sample units), and
4. errors not distributed normally.

The first problem is overcome by considering a variety of linear and nonlinear models. Nonrandom patterns in plots of residuals versus predicted volumes show when a model fits the data poorly. With multiple linear regression this also has led to the use of incremental or stepwise inclusion of variables or the evaluation of all possible combinations of independent variables. The second problem may be minimized by weighting observations or by using transformations (for example, taking logarithms) to equalize variance of errors (Cunia 1964). The third is difficult to deal with and is the subject of recent procedures like the two-stage approaches developed by econometricians (Judge et al. 1985). For data with correlated errors, estimates of model parameters are unbiased, but hypothesis tests may be affected by inflated estimates of variance. The fourth problem is not serious since the central limit theorem tends to make the distribution of means normal.

Stepwise regression procedures allow the analysis to determine which independent variables should be used to predict the response. Since these procedures are based on the data, they may not always achieve the same results and may not find the best relationship. It is essential to understand what rules are being used in these statistical selection procedures. An alternative is to evaluate a set of candidate relationships which have been selected based on the prior work of other researchers or on the basis of theoretical expectations.

The strength of regression relationships is often summarized with only a few statistics, most notably the standard error of estimate (square root of error variance) and the coefficient of determination (R-squared). Although popular and almost universally reported, the R-squared statistic is only one of several that should be used in evaluating the performance of a regression relationship. Remember, for transformed models both the R-squared coefficient and standard error of estimate are calculated on the basis of the transformed dependent variable, not the original variable which is normally of more interest. For nonlinear regressions the R-squared statistic has no theoretical interpretation as it does in linear regression, although an empirical value can be calculated. When sample size is small, the adjusted R-squared statistic is preferred since it is adjusted for sample size.

The goodness-of-fit for a relationship should be evaluated using a number of statistics. Scatterplots can provide a good visual impression of basic relationships in the data. The correlation matrix may also assist in evaluating interrelations among all the variables. An R-squared coefficient greater than 0.9 and a standard error of estimate less than 20% of the mean can be expected for a well defined tree volume equation. A plot of residuals (observed minus predicted values for each observation) versus predicted values allows identification of several potential problems including outliers, unequal variance, poor model specification, or indications of potential bias in different parts of the range of predicted values. Histograms of residuals or other tests can be used to determine whether the errors are normally distributed. Correlated errors are more difficult to detect; a careful evaluation of sampling procedures may identify likely sources of correlation.

The contribution of individual independent variables should be evaluated by testing the hypothesis that the parameter is zero. In simple linear regression, the slope term may be tested using a t-test to see whether it is really different from zero. Concluding that the slope is not different from zero is the same as saying there is no relation between the independent and dependent variables. The same test may be generalized to multiple regression where, for example, the slope coefficient for form class might be effectively zero if the diameter and height are already included in the predictive model.

While assumptions are particularly important for hypothesis testing, the ultimate test of a regression relationship may lie in its use for prediction. Neter and Wasserman (1974) suggest the evaluation of prediction bias by evaluating performance of the relation when used to predict values from an independent data set not used in the estimation of parameters. Verbyla (1986) showed that prediction bias can be substantial when many independent variables are used with small sample sizes.

Finally, a report is desirable summarizing characteristics of alternative models and the selection of the final model. Other desirable information should be listed, including:

1. author and date,
2. number, species, and size range of source data,
3. location or area,
4. description of sample tree selection procedures,
5. measurement specifications for variables, and
6. regression statistics
   a) error variance,
   b) adjusted R-squared value,
   c) t statistics for coefficient estimates,
   d) residuals plots (residual vs predicted), and
   e) error variance estimated for independent data.

4. Using Tree Volume Relations

4.1 Measurement Specifications and their Impact on Usage

When using existing volume relationships it is essential to follow the same measurement standards as those on which the relationship is based. The most obvious error is to use a merchantable height with a volume relation based on total height. Less obvious, and less serious, is to use a diameter measurement on the uphill side of the tree if the relation is based on diameters measured at midslope.

Since height is usually only measured for a subsample of trees, several approaches have been used to improve on local volume relations by predicting height as a function of diameter. Any of several sigmoidal growth models might be used; the exponential models offer one example (Avery and Burkhart 1983):

$$H = \exp(a + b / D) \qquad [8]$$

This model may be substituted for height in the nonlinear power model (the second standard tree volume model, equation [3], described earlier) to give:

$$V = a \times D^b \times \exp(c + d / D) \qquad [9]$$

This model may be fitted, as a local volume function, to data consisting of diameters and volumes with the expectation that knowing the form of the height development curve in relation to diameter may aid in predicting volume. Alternatively, if other data are available with some height measurements, then separate height prediction equations can be developed and applied to trees for which only diameter was measured. Measured diameters and predicted heights can then be used with a standard volume equation.

4.2 Extrapolation

Predictions of volume beyond the range of data used to establish the relationship is risky and should be avoided if possible. If necessary, the ability of the model to predict reasonable values should be considered. For example, a local volume relation based on the square of diameter is likely to overestimate volume for extremely large trees. An exponential model, such as $V = \exp(b_0 + b_1 / D)$, would be better suited for occasionally extrapolating to extremely large trees.

4.3 Adjustment of Existing Relationships

If existing tree volume relationships are used in new applications where there is any question about their appropriateness, adjustments may be made using a ratio/regression sampling procedure. In the simplest case the ratio estimate would be:

$$VADJ = R \times V \quad [10]$$

where R = ratio of average measured volume to average predicted volume for a sample of trees, and

VADJ = predicted volume based on proportion adjustment.

5. Conclusion

Success in evaluating tree volume estimation relationships requires careful planning and execution coordinated among a number of activities, including (1) underlying theory, (2) population definition, (3) sampling procedures, (4) measurement methods, and (5) curve fitting techniques. Extrapolation beyond the limits of these activities should be avoided.

6. References

Avery, T.E.; Burkhart, H.E. 1983. Forest measurements. 3rd ed. McGraw-Hill, New York, N.Y. 331 p.

Clutter, J.L.; Fortson, J.C.; Pienaar, L.V.; Brister, G.H.; Bailey, R.L. 1983. Timber management: a quantitative approach. Wiley & Sons, New York, N.Y. 333 p.

Cunia, T. 1964. Weighted least squares method and construction of volume tables. For. Sci. 10(2):180-191.

Demaerschalk, J.P.; Kozak, A. 1974. Suggestion and criteria for more effective regression sampling. Can. J. For. Res. 4:341-348.

Demaerschalk, J.P.; Kozak, A. 1975. Suggestion and criteria for more effective regression sampling 2. Can. J. For. Res. 5:496-497.

Gove, J.H.; Barnett, J.P.; Gregoire, T.G. 1982. When is *n* sufficiently large for regression estimation. J. Environ. Manage. 15:229-237.

Husch, B.; Miller, C.I.; Beers, T.W. 1982. Forest mensuration. 3rd ed. John Wiley & Sons, Inc., New York, N.Y. 402 p.

Judge, G.G.; Hill, R.; Griffiths, W.; Lutkepohl, H.; Lee, T. 1985. Introduction to the theory and practice of econometrics. 2nd ed. Wiley, New York, N.Y. 839 p.

Marshall, P.L.; Demaerschalk, J.P. 1986. A strategy for efficient sample selection in simple linear regression problems with unequal per unit sampling costs. For. Chron. 62:16-19.

Neter, J.; Wasserman, W. 1974. Applied linear statistical models. Richard D. Irwin, Inc., Homewood, Il. 842 p.

Schumacher, F.X.; Hall, F.S. 1933. Logarithmic expression of timber-tree volume. J. Agric. Res. 47:719-731.

Spurr, S.H. 1952. Forest inventory. Ronald Press, New York, N.Y. 476 p.

Titus, S.J.; Morgan, D.J. 1985. Tree height: Can large-scale photo measurements be more accurate than field measurements? For. Chron. 61:214-217.

West, P.W.; Ratkowsky, D.A.; Davis, A.W. 1984. Problems of hypothesis testing of regressions with multiple measurements from individual sampling units. For. Ecol. Manage. 7:207-224.

Verbyla, D. 1986. Potential prediction bias in regression and discriminant analysis. Can. J. For. Res. 16:1255-1257.

Gouvernement du Canada
Government of Canada

Service canadien des forêts
Canadian Forestry Service

Théorie et méthodes d'estimation du volume individuel des arbres

S.J. Titus et D.J. Morgan

Rapport d'information PI-X-85F
Institut forestier national de Petawawa et Comité canadien de l'inventaire forestier

INSTITUT FORESTIER NATIONAL DE PETAWAWA

Le mandat de l'Institut forestier national de Petawawa, comme celui des autres établissements du Service canadien des forêts, est de promouvoir une meilleure gestion et une utilisation plus rationnelle des ressources forestières du Canada, pour le bien économique et social de tous les Canadiens. Les objectifs des activités des programmes menés à l'Institut appuient ce mandat à travers la découverte, le développement, la démonstration, l'application et le transfert des innovations. En tant qu'institut national, il doit s'attacher à des problèmes qui débordent le cadre régional ou qui nécessitent des compétences particulières de même qu'un équipement non disponible aux installations régionales du Service canadien des forêts. La plupart du temps, ses recherches sont effectuées en étroite collaboration avec le personnel des centres régionaux, des services forestiers des provinces et de l'industrie forestière.

Les projets de recherche et les services techniques de l'Institut sont regroupés sous cinq principales activités:

GÉNÉTIQUE FORESTIÈRE ET BIOTECHNOLOGIE — Ce programme encadre des projets sur la génétique forestière, la microbiologie, la micropropagation, la génétique moléculaire et la recherche sur les semences. Il comprend également les services à la clientèle et la banque de semences du Centre national de semences forestières. Ce centre existe depuis longtemps et est lié à plusieurs organismes internationaux.

SYSTÈMES D'AMÉNAGEMENT FORESTIER — Ce programme intègre en recherche et développement, des projets sur les incendies de forêt, la télédétection, la météorologie, la modélisation, la croissance et la récolte, et le RIMA. Il permet ainsi l'élaboration et la démonstration de systèmes d'aménagement forestier.

STATISTIQUES NATIONALES SUR LES RESSOURCES FORESTIÈRES — Ce programme fournit l'information biologique, technique et socio-économique sur les ressources forestières du Canada. Il vise le développement des bases de données, telle que FIDSINFOBASE, et l'établissement de nouveaux logiciels et bases de données pour aider à l'élaboration des politiques forestières. Le Programme d'inventaire forestier recueille l'information sur les forêts au niveau national, maintient le Système sur les ressources forestières canadiennes, et prépare l'inventaire des forêts du Canada.

COMMUNICATIONS — Ce programme regroupe les activités de la bibliothèque et des projets touchant la sensibilisation du public, l'information, la rédaction-révision et les publications. L'Institut reçoit plus de 20 000 visiteurs chaque année. Le Centre d'accueil, des visites autoguidées et un programme éducatif complet sont ouverts à tous. C'est à l'IFNP que l'on trouve l'entrepôt et le centre de distribution nationals de toutes les publications scientifiques du SCF.

LA FORÊT EXPÉRIMENTALE — Outre des peuplements naturels où l'on applique divers traitements dans le cadre de projets de recherche en sylviculture, la Forêt de Petawawa, mesurant 98 km^2, comprend de vastes superficies de plantations âgées de plus de 60 ans. Ces plantations expérimentales permettent d'obtenir des données sur la croissance et la récolte à la suite d'expériences de culture. Elles fournissent du matériel génétique dont le lignage est répertorié, et se révèlent de plus en plus utiles pour les études sur la micropropagation et la génétique moléculaire. La forêt permet également de mettre à l'essai des stratégies d'aménagement forestier à court et à long termes.

THÉORIE ET MÉTHODES D'ESTIMATION DU VOLUME INDIVIDUEL DES ARBRES

S.J. Titus, Département des sciences forestières, Université de l'Alberta
D.J. Morgan, Direction de l'aménagement forestier, Service des forêts,
ministère des Forêts, des Terres et de la Faune de l'Alberta

Rapport d'information PI-X-85F
Publication conjointe n° 1
Institut forestier national de Petawawa
et Comité canadien de l'inventaire forestier
1988

©Ministre des Approvisionnements et Services Canada, 1988
N° de catalogue Fo46-11/85-1988
ISBN 0-662-56318-2
ISSN 0228-0736

Il est possible d'obtenir sans frais des exemplaires de cette publication auprès du/de:

Centre d'information technique et de distribution
Institut forestier national de Petawawa
Chalk River (Ontario)
K0J 1J0

N° de téléphone: 613-589-2880

Des copies ou microfiches de cette publication sont en vente chez:

Micromédia Ltée
Place du Portage
165, rue Hôtel-de-Ville
Hull (Québec)
J8X 3X2

Comité canadien de l'inventaire forestier

Le Comité canadien de l'inventaire forestier (CCIF) est un comité national dont le mandat est d'examiner tout ce qui touche à l'inventaire forestier au Canada, qu'il s'agisse de la recherche ou des pratiques établies au pays pour la tenue des relevés, la mise à jour des inventaires et la collecte d'autres données sur l'aménagement forestier. Ce mandat s'étend aux méthodes administratives, aux normes de mesure, à la terminologie ainsi qu'à l'acquisition, au traitement et à la diffusion des données.

Les membres votants de ce comité sont les directeurs de l'inventaire forestier au sein des organismes de foresterie provinciaux et territoriaux et du Service canadien des forêts ainsi qu'un membre indépendant. D'autres membres non votants apportent au comité leur compétence dans divers aspects spécialisés de l'inventaire forestier.

Table des matières

		Page
Résumé		*iv*
1.	Introduction	1
2.	Théorie des relations entre le volume de l'arbre et d'autre variables	1
	2.1 Variables	1
	2.2 Relations entre le volume de l'arbre et d'autres variables	1
3.	Méthodes d'évaluation des relations entre le volume de l'arbre et d'autres variables	2
	3.1 Caractéristiques de la population	3
	3.2 Méthodes d'échantillonnage	3
	3.3 Techniques de mesure	3
	3.4 Ajustement de courbe	5
4.	Utilisation des relations entre le volume de l'arbre et d'autres variables	7
	4.1 Caractéristiques du mesurage et incidences concrètes	7
	4.2 Extrapolation	7
	4.3 Ajustement des relations existantes	7
5.	Conclusion	8
6.	Bibliographie	8

Résumé

Ce rapport présente un examen de la théorie et des méthodes d'estimation du volume individuel des arbres. Il traite notamment de lignes directrices relatives 1) à la définition de la population, 2) aux méthodes d'échantillonnage, 3) aux techniques de mesure et 4) aux techniques d'ajustement de courbe.

1. Introduction

Le volume est une variable fondamentale pour décrire les dimensions et le potentiel de matière ligneuse des différents arbres et établir des estimations du volume des peuplements. L'objectif du présent document est de résumer les connaissances acquises jusqu'ici sur les principales relations utilisées pour prédire le volume du fût d'un arbre. Nous accordons une attention particulière aux méthodes servant à évaluer ces relations, de même qu'à des considérations d'ordre dendrométrique et statistique. Notre objectif est de présenter une démarche aux praticiens de la foresterie qui doivent mettre au point ou évaluer des relations pour estimer le volume des arbres. Il s'agit ici d'une version revue et augmentée d'un avant-projet de liste de contrôle des méthodes de mesure, établi précédemment par le sous-comité de dendrométrie du Comité canadien de l'inventaire forestier. Nous avons exclu de la discussion un certain nombre de sujets secondaires connexes, dont les défauts, la qualité, le potentiel de commercialisation, les rapports avec le défilement et la définition du fût chez les arbres déliquescents (généralement des feuillus).

2. Théorie des relations entre le volume de l'arbre et d'autres variables

Le volume d'un arbre peut être obtenu à l'aide de plusieurs techniques de mesure directe, ou indirectement estimé à partir de la relation entre le volume et d'autres caractéristiques de l'arbre. Le mesurage direct consiste, la plupart du temps, à couper l'arbre en sections et à additionner les volumes estimés des sections pour obtenir le volume total de l'arbre. Il est quelquefois possible de mesurer le volume des arbres sur pied à l'aide d'instruments comme un dendromètre optique pour obtenir des mesures du diamètre et de la hauteur à différents endroits du fût.

La méthode qui consiste à mesurer le déplacement d'un liquide, tout en étant idéale pour estimer le volume de sections d'arbre, est rarement utilisée dans les faits. On calcule le plus souvent les volumes des sections en appliquant une formule comme celle de Smalian qui suppose que les sections sont des troncs de paraboloïdes. Les approches directes étant souvent coûteuses et fastidieuses, le volume des arbres sur pied est souvent estimé à partir de la relation entre le volume et des caractéristiques de l'arbre faciles à mesurer comme le diamètre, la hauteur et la forme.

2.1 Variables

La variable dépendante est le volume brut du fût de l'arbre en mètres cubes, à l'intérieur de l'écorce, du niveau du sol jusqu'à l'extrémité de la flèche de l'arbre. Un certain nombre de variables indépendantes sont liées au volume de l'arbre, y compris la hauteur de l'arbre (m) et le diamètre (cm) à 1,3 m de hauteur. De plus, on peut «fixer» d'autres variables comme l'essence ou la région en élaborant des équations distinctes pour chaque combinaison afin de réduire la variabilité de la population. D'autres variables comme le rapport hauteur cime/hauteur totale et une mesure de la forme de l'arbre peuvent parfois être prises en compte. Toutefois, lorsqu'on se sert de relations pour estimer le volume des arbres dans le cadre de travaux réguliers d'inventaire, il peut être nécessaire de limiter les variables indépendantes à celles habituellement mesurées et notées au cours d'un inventaire.

2.2 Relations entre le volume de l'arbre et d'autres variables

Pour estimer le volume d'un arbre, il est utile de se représenter le fût (partie du tronc située au-dessus de la souche et sous le houppier) comme un cône. On obtient le volume d'un cône en intégrant

la surface de la section transversale, qui varie selon la décroissance (défilement) linéaire, de la base au sommet du cône, ce qui donne la formule suivante:

$$V = (H / 3) \times \pi (D / 200)^2 \quad [1]$$

où le volume (V), la hauteur (H) et le diamètre (D) sont respectivement exprimés en mètres cubes, en mètres et en centimètres. Ce modèle géométrique simple du volume de l'arbre sert de fondement à la plupart des relations valables utilisées pour estimer le volume des arbres sur pied. On obtient des variations de ce modèle en généralisant le modèle de défilement sous-jacent pour prendre en compte d'autres formes de défilement dont le paraboloïde et le néloïde.

Vous trouverez ci-dessous trois exemples de relations, d'une part pour le volume standard et d'autre part pour le volume local. D'autres formes de modèles ont fait l'objet de recherches et de rapports (Cluster et collab., 1983; Husch et collab., 1982; Spurr, 1952). Les termes «modèle», «formule» et «équation» sont synonymes dans le texte.

Trois relations ont été utilisées pour estimer le volume standard et toutes reposent sur des variations de la formule du volume d'un cône:

$$V = a + b \times D^2 \times H \quad [2]$$
$$V = a \times D^b \times H^c \quad [3]$$
$$\log V = a + b \times \log D + c \log H \quad [4]$$

où a, b, c et d sont des paramètres des modèles et log est le logarithme (traditionnellement de base 10). Le troisième modèle (Schumacher et Hall, 1933) est la forme linéarisée du deuxième modèle qui est lui-même une légère généralisation du premier (Spurr, 1952).

Les relations qui concernent le volume local n'utilisent généralement que le dhp pour estimer le volume sont par conséquent sujettes à des erreurs plus grandes:

$$V = a + b \times D^2 \quad [5]$$
$$V = a \times D^b \quad [6]$$
$$\log V = a + b \log D \quad [7]$$

Le deuxième modèle est une généralisation du premier, permettant d'estimer la puissnce du diamètre; le troisième modèle est la forme logarithmique du second. En télédétection, on désigne quelquefois comme une relation de volume local l'estimation du volume à partir de la hauteur mesurée sur une photographie aérienne.

3. Méthodes d'évaluation des relations entre le volume de l'arbre et d'autres variables

La présente section est divisée en quatre parties portant respectivement sur (1) les caractéristiques de la population, (2) les méthodes d'échantillonnage, (3) les techniques de mesure et (4) l'ajustement de courbe. Même si ces sujets sont présentés séparément, les uns à la suite des autres, ils sont tous reliés et chaque partie traite nécessairement de certains aspects des autres parties en raison de leur importance et de leur impact. Ainsi, les méthodes d'échantillonnage et les techniques de mesure sont affectées par les hypothèses et les exigences de l'ajustement de courbe.

3.1 Caractéristiques de la population

Il faut établir soigneusement les caractéristiques de la population afin de connaître les limites de l'inférence. Il faut souvent limiter le champ d'application d'une relation à une espèce et à une région géographique données. Il faut non seulement préciser l'essence et la région en question, mais également les caractéristiques des différents arbres. Faut-il inclure les arbres fourchus ou étêtés? Faudrait-il limiter la population aux arbres n'ayant pas encore atteint l'âge d'exploitabilité caractéristique de cette essence? Les réponses à ces questions différeront dans la mesure où les prévisions portent sur du matériel sur pied relatif cultivé dans des plantations ou sur la population naturelle existante.

3.2 Méthodes d'échantillonnage

Le prélèvement d'échantillons pour établir une relation (une régression) est souvent très différent de l'échantillonnage à des fins d'inventaire. Il est important, cette fois, d'échantillonner sur l'étendue entière des variables indépendantes et dépendantes, ce qui peut exiger l'utilisation d'une méthode d'échantillonnage moins aléatoire et plus sélective et représentative.

L'idéal serait de sélectionner des arbres-échantillons dans des endroits choisis au hasard dans l'ensemble de l'aire d'extension naturelle de la population, mais plusieurs arbres sont souvent mesurés au même endroit par souci d'économie. Toutefois, lorsque le nombre d'arbres choisis dans un même endroit devient trop élevé par rapport au nombre total d'endroits retenus, les conditions échantillonnées peuvent ne pas être représentatives de l'ensemble d'une population puisque les arbres poussant les uns près des autres ont tendance à se ressembler. West et collab. (1984) ont démontré que de multiples observations dans une même unité d'échantillonnage, comme une parcelle, peuvent produire des estimations biaisées des variances lors de la vérification des hypothèses des modèles de régression.

Il faut un nombre raisonnable d'arbres-échantillons, mais le nombre exact dépend de la variabilité de la population, de la force des rapports, de l'amplitude des erreurs acceptables et des contraintes de coûts (Demaerschalk et Kozak, 1974 et 1975; Gove et collab., 1982; Marshall et Demaerschalk, 1986). Un minimum de 30 arbres-échantillons donnera une indication de la nature fondamentale de la relation, mais il peut être souhaitable de disposer d'un échantillon de 100 arbres ou plus, selon la variabilité et l'éventail des dimensions des arbres. De plus, la répartition des unités échantillonnées dans la gamme des valeurs des variables indépendantes et dépendantes devrait être uniforme uniquement si la variance des erreurs est la même pour les différentes valeurs des variables indépendantes. Si la variance de l'erreur augemente avec l'accroissement des valeurs de la variable dépendante, comme c'est habituellement le cas avec le volume des arbres, la répartition devrait alors être faite proportionnellement à la variance et les probabilités d'échantillonner des arbres de dimensions supérieures devraient être plus grandes (comme c'est le cas avec l'échantillonnage par placettes circulaires à rayon variable ou «par balayage sous angle constant»). Des données de ce type doivent être soumises à une analyse de «pondération »(voir paragraphes suivants). Il n'est pas facile de déterminer la taille adéquate d'un échantillon. Il faut se rappeler qu'un échantillon plus gros ne sera pas automatiquement meilleur; la qualité et la répartition des données sont également des facteurs importants.

3.3 Techniques de mesure

Le mesurage des arbres-échantillons exige le recours à des techniques uniformes et la réduction au minium des erreurs puisque l'analyse ultérieure des données destinée à évaluer les relations repose sur l'hypothèse que les erreurs de mesurage sont négligeables.

Le diamètre des arbres est généralement mesuré à l'aide d'un compas forestier ou d'un ruban diamétrique. Si l'on utilise un compas forestier, il faudra prendre deux mesures, à angle droit l'une de l'autre, la moyenne géométrique servant de diamètre moyen.

La hauteur devrait être établie à partir du mesurage des arbres abattus ou de mesures précises des arbres sur pied afin de réduire les erreurs au minimum. Les méthodes d'inventaire habituelles comme l'utilisation d'un clinomètre ou des distances mesurées au nombre de foulées ou au ruban ne sont en général par suffisamment précises (Titus et Morgan, 1985).

Si les arbres sont abattus et coupés en sections pour en déterminer le volume, la coupe la plus basse devrait être effectuée au niveau du sol ou à la hauteur de la souche (0,3 m), ou aux deux endroits, suivie d'une autre coupe à hauteur de poitrine (1,3 m). Avant l'abattage, il faut mesurer et marquer sur chaque arbre la hauteur de poitrine et la hauteur de la souche. Il faut mesurer le diamètre sur écorce à ces endroits pour s'assurer que les dommages causés par l'abattage n'influent pas sur les dimensions. Le marquage assure également que ces points critiques sont placés de façon uniforme. Les sections coupées à plus de 1,3 m de hauteur devraient avoir moins de 2,0 m afin de réduire au minimum les possibilités d'erreurs d'estimation du volume attribuables au défilement. L'épaisseur de l'écorce peut être mesurée en plus du diamètre, ou encore on peut mesurer le diamètre sur écorce (ib) et le diamètre sous écorce (ob). Les distorsions localisées de la tige comme les broussins ou les couronnes de branches peuvent être évitées en modifiant, au besoin, la longueur des sections.

En pratique, les méthodes varient, allant de la sélection de courtes sections de longueur déterminée (par exemple 1,0 m) ou d'un nombre déterminé de sections (par exemple 10 sections) de chaque arbre, peu importe sa hauteur. Les sections courtes de longueur déterminée permettent d'obtenir de meilleures estimations du volume, notamment lorsque les variations extrêmes du défilement sont courantes, mais cette façon de procéder peut être coûteuse lorsque les arbres sont grands. La deuxième méthode permet de réduire les coûts, mais peut donner des estimations très inexactes du volume des sections (un arbre de 30 m, à raison de 10 sections par arbre, donnera des sections d'une longueur de près de 3 m). La sélection d'un compromis approprié exige un équilibrage des coûts avec le degré d'exactitude désiré. Nous suggérons les normes de mesurage suivants:

diamètre des sections	-0,1 cm
longueur des sections	-0,01 m
volume	-0,001 m^3

Il est recommandé d'assimiler les formes de sections aux types dendrométriques suivants pour estimer les volumes des différentes sections:

souche	- cylindre de tronc de néloïde
partie médiane	- tronc de paraboloïde
partie terminale	- cône

Il faut utiliser les mêmes types dendrométriques lorsqu'on doit faire des interpolations à l'intérieur de sections. Il est à noter qu'à mesure que la longueur d'une section diminue, la différence de diamètre (surface) à chaque extrémité décroît et les différences de volume calculé avec les différentes équations deviennent négligeables. Inversement, plus la longueur des sections augmente, plus les différences calculées à l'aide des différentes formules s'élargissent et plus les hypothèses de forme logarithmiques prennent de l'importance. Notre recommandation d'utiliser la formule de Smalian pour les sections médianes d'un arbre repose sur l'hypothèse additionnelle que les sections sont courtes puisque cette méthode surestime le volume des sections longues (Husch et collab., 1982).

Il faut passer au crible toutes les données avant d'effectuer l'analyse statistique afin de détecter les erreurs et les «observations extrêmes». Les erreurs devront être corrigées dans la mesure du possible; les observations extrêmes devront être mises de côté si l'on peut établir qu'une erreur

évidente a été commise et ne peut être corrigée. Les éléments de données ne devraient par être biffés uniquement parce qu'ils semblent inhabituels. Des sommaires graphiques des données sont utiles pour identifier les observations extrêmes et peuvent également aider à effectuer des évaluations préliminaires de la nature des relations entre le diamètre, la hauteur et le volume.

3.4 Ajustement de courbe

L'ajustement de courbe, ou l'analyse de régression, peut être linéaire ou non linéaire, selon la forme du modèle utilisé. Les modèles linéaires sont davantage utilisés pour un certain nombre de raisons, notamment parce qu'ils reposent sur une théorie statistique mieux élaborée et que les méthodes d'ajustement de courbe sont plus faciles. Avec l'accroissement des puissances de calcul disponibles, les modèles non linéaires ont gagné en popularité car ils permettent d'obtenir des formes de modèle qui correspondent aux attentes théoriques. L'équation non linéaire peut également être considérée comme une généralisation du modèle linéaire qui permet l'ajustement tant du modèle linéaire que du modèle non linéaire.

Les modèles de régression permettent les erreurs résiduelles, selon l'hypothèse que la moyenne à long terme des erreurs est nulle, que les erreurs sont indépendantes et que la variance des erreurs est constante à tous les niveaux des variables indépendantes. Dans un régression linéaire, on suppose de plus que les erreurs suivent une distribution normale.

Ces hypothèses sont souvent laissées de côté, ce qui entraîne l'utilisation de données et de méthodes d'analyse inadéquates. Plusieurs genres de problèmes peuvent survenir, les plus courants étant les suivants:

1. mauvais ajustement du modèle (souvent dit «biaisé»)
2. variance inégale des erreurs (augmentant généralement en même temps que la variable indépendante)
3. erreurs corrélées (mesures multiples des mêmes unités d'échantillonnage) et
4. erreurs non normalement distribuées

On pourra surmonter le premier problème en envisageant une variété de modèles linéaires et non linéaires. Des structures non aléatoires dans la représentation graphique des résidus en fonction des volumes prévus sont le signe d'un mauvais aujustement du modèle aux données. Ce problème incite également à recourir à la régression linéaire multiple, qui permet l'inclusion de variables par la méthode pas à pas ou l'évaluation de toutes les combinaisons possibles de variables indépendantes. Le deuxième problème peut être minimisé par la pondération des observations ou à l'aide de transformations (par exemple en logarithmes) pour égaliser la variance des erreurs (Cunia, 1964). Le troisième est plus difficile à régler et est abordé par des méthodes récentes comme l'approche en deux étapes élaborée par des économétriciens (Judge et collab., 1985). Dans le cas des données à erreurs corrélées, les estimations des paramètres du modèle sont non biaisées, mais les tests d'hypothèses peuvent être perturbés par des estimations exagérées de la variance. Le quatrième problème n'est pas grave, car en vertu du théorème de la limite centrale, la distribution des moyennes tend vers la normale.

Les méthodes de régression pas à pas permettent d'effectuer une analyse pour déterminer quelles variables indépendantes utiliser pour prédire les résultats. Ces méthodes étant basées sur les données, elles peuvent ne pas toujours donner les mêmes résultats et ne pas permettre d'établir la meilleure relation. Il est essentiel de bien comprendre quelles sont les règles utilisées dans ces méthodes statistiques de sélection. Une solution de rechange serait d'évaluer un ensemble de relations possibles, choisies d'après des travaux antérieurs d'autres chercheurs ou en fonction des attentes théoriques.

La force des relations de régression n'est souvent résumée que par quelques statistiques, le plus souvent l'erreur-type d'estimation (racine carrée de la variance de l'erreur) et le coefficient de détermination (R carré). Ce coefficient de détermination R carré, tout en étant populaire et universellement reconnu, n'est que l'un des nombreux critères qui devraient être utilisés pour évaluer la performance d'une relation de régression. Ne jamais oublier que le coefficient de détermination et l'erreur-type d'estimation des modèles transformés sont toujours calculés à partir de la variable dépendante transformée et non pas de la variable d'origine qui est habituellement plus pertinente. Le R carré n'a aucune interprétation théorique dans le cas des régressions non linéaires, au contraire des régressions linéaires, bien qu'une valeur empirique puisse être calculée. Lorsque l'échantillon est petit, le R carré ajusté est préférable puisqu'il est redressé pour tenir compte de la taille de l'échantillon.

La validité de l'ajustement d'une relation peut être évaluée à l'aide d'un certain nombre de statistiques. Les graphiques de dispersion peuvent permettre de se faire une assez bonne idée des relations fondamentales contenues dans les données. La matrice des corrélations peut également nous aider à évaluer les relations mutuelles entre les variables. Il faut s'attendre à ce qu'une équation bien définie de cubage des arbres ait un coefficient de détermination supérieur à 0,9 et une erreur-type d'estimation inférieure à 20 % de la moyenne. Un graphique des résidus (valeurs observées moins valeurs estimées pour chaque observation) par rapport aux valeurs estimées permet de déceler plusieurs problèmes éventuels, notamment des valeurs aberrantes, une variance inégale, une spécification inadéquate du modèle, ou donne des indices sur les biais éventuels dans les différentes parties de l'étendue de variation des valeurs estimées. Des histogrammes des résidus ou d'autres tests peuvent servir à déterminer si les erreurs sont normalement distribuées. Il est plus difficile de détecter les erreurs corrélées; une évaluation approfondie des méthodes d'échantillonnage peut permettre de repérer les sources éventuelles de corrélation.

La contribution de chacune des variables indépendantes devrait être évaluée en vérifiant l'hypothèse que le paramètre est nul. La pente d'une régression linéaire simple peut être vérifiée à l'aide d'un test pour voir si elle est vraiment différente de zéro. Arriver à la conclusion que la pente n'est pas différente de zéro équivaut à dire qu'il n'existe aucune relation entre la variable indépendante et la variable dépendante. Ce même test peut être généralisé à une régression multiple dans laquelle, par exemple, le coefficient de pente pour les classes de forme peut effectivement être nul si le diamètre et la hauteur sont déjà inclus dans le modèle prévisionnel.

Bien que les postulats soient particulièrement importants pour la vérification des hypothèses, une relation établie par régression sera évaluée, en fin de compte, en termes de sa capacité de prévision. Neter et Wasserman (1974) suggèrent d'établir l'erreur systématique de prévision en évaluant la performance de la relation lorsqu'elle est utilisée pour estimer les valeurs d'un ensemble de données indépendantes qui n'ont pas été employées pour estimer les paramètres. Verbyla (1986) a montré que l'erreur systématique de prévision peut être importante lorsque de nombreuses variables indépendantes sont utilisées avec de petits échantillons.

Pour terminer, disons qu'il est souhaitable de rédiger un rapport résumant les différents modèles possibles et expliquant la sélection du modèle définitif. Il faudrait énumérer d'autres renseignements utiles notamment:

1. le nom de l'auteur et la date
2. le nombre, les essences et l'étendue de variation des données de base
3. l'endroit ou la région
4. une description des méthodes de sélection des arbres-échantillons
5. les caractéristiques du mesurage des variables et
6. les statistiques de la régression
   a) variance de l'erreur
   b) R carré ajusté

c) statistique t pour les estimations des coefficients
d) diagrammes des résidus (résidus par rapport aux prévisions) et
e) variance de l'erreur estimée pour des données indépendantes

4. Utilisation des relations entre le volume de l'arbre et d'autres variables

4.1 Caractéristiques du mesurage et incidences concrètes

Lors de l'emploi des relations de cubage, il est essentiel de respecter les mêmes normes que celles ayant présidé à l'élaboration de la relation. L'erreur la plus courante est d'utiliser la hauteur marchande dans une équation de cubage basée sur la hauteur totale. Une erreur moins évidente et moins grave consiste à utiliser une mesure du diamètre prise en se tenant du côté de l'arbre situé vers le haut de la pente si la relation est basée sur des diamètres mesurés au milieu de la pente.

La hauteur n'étant généralement mesurée que sur un sous-échantillon d'arbres, plusieurs approches ont été utilisées pour améliorer les relations de cubage local en estimant la hauteur en fonction du diamètre. L'un des nombreux modèles de croissance sigmoïde peut être utilisé; les modèles exponentiels en sont un exemple (Avery et Burkhart, 1983):

$$H = \exp(a + b/D) \qquad [8]$$

Ce modèle peut être substitué à la hauteur dans le modèle logarithmique non linéaire (le deuxième modèle de cubage standard, soit l'équation (3) décrite précédemment), ce qui donne:

$$V = a \times D^b \times \exp(c + d/D) \qquad [9]$$

On peut ajuster ce modèle, en tant que fonction de cubage local, à des données constituées de diamètres et de volumes, en misant sur le fait que la connaissance de la forme de la courbe du développement en hauteur par rapport au diamètre nous aidera à prévoir le volume. Si l'on dispose d'autres données, dont certaines sur la hauteur, on pourra alors élaborer des équations distinctes pour estimer la hauteur et les appliquer à des arbres dont seul le diamètre a été mesuré. Les diamètres mesurés et les hauteurs estimées peuvent alors être employés dans une équation de cubage standard.

4.2 Extrapolation

Il est hasardeux et, dans la mesure du possible, contre-indiqué d'estimer un volume au-delà de l'étendue de variation des données utilisées pour établir la relation. Il faudra, au besoin, examiner la capacité du modèle à prévoir des valeurs raisonnables. Ainsi, une équation de cubage local basée sur le carré des diamètres est susceptible de surestimer le volume des arbres extrêmement gros. Un modèle exponentiel comme $V = \exp(b_0 + b_1/D)$ serait plus indiqué pour obtenir occasionnellement par extrapolation le volume de très gros arbres.

4.3 Ajustement des relations existantes

Si des équations de cubage existantes servent à de nouvelles applications, mais que leur pertinence n'est pas certaine, on peut procéder à des ajustement à l'aide d'une méthode de correction par quotient. L'estimation ajustée, dans le plus simple des cas, serait la suivante:

$$V_{ADJ} = R \times V \qquad [10]$$

où R = rapport entre le volume mesuré moyen et le volume estimé moyen pour un échantillon d'arbres, et

V_{ADJ} = volume estimé après correction par quotient.

5. Conclusion

Le succès de l'évaluation des relations d'estimation du volume des arbres exige qu'on apporte le plus grand soin à la planification et à la réalisation coordonnées de plusieurs activités: (1) étude de la théorie sous-jacents, (2) définition de la population, (3) élaboration des méthodes d'échantillonnage, (4) établissement des méthodes de mesure et (5) choix des méthodes d'ajustement de courbe. Il faudrait éviter d'extrapoler au-delà de ces activités.

6. Bibliographie

Avery, T.E.; Burkhart, H.E. 1983. Forest measurements. 3rd ed. McGraw-Hill, New York, N.Y. 331 p.

Clutter, J.L., Fortson, J.C., Pienaar, L.V., Brister, G.H.; Bailey, R.L. 1983. Timber management: a quantitative approach. Wiley & Sons, New York, N.Y. 333 p.

Cunia, T. 1964. Weighted least squares method and construction of volume tables. For. Sci. 10(2), 180-191.

Demaerschalk, J.P.; Kozak, A. 1974. Suggestion and criteria for more effective regression sampling. Journ. can. rech. for. 4:341-348.

Demaerschalk, J.P.; Kozak, A. 1975. Suggestion and criteria for more effective regression sampling 2. Journ. can. rech. for. 5: 496-497.

Gove, J.H., Barnett, J.P.; Gregoire, T.G. 1982. When is *n* sufficiently large for regression estimation. J. Environ. Manage. 15: 229-237.

Husch, B.; Miller, C.I.; Beers, T.W. 1982. Forest mensuration. 3rd ed. John Wiley & Sons, Inc., New York, N.Y. 402 p.

Judge, G.G.; Hill, R.; Griffiths, W.; Lutkepohl, H.; Lee, T. 1985. Introduction to the theory and practice of econometrics. 2nd ed. Wiley, New York, N.Y. 839 p.

Marshall, P.L.; Demaerschalk, J.P. 1986. A strategy for efficient sample selection in simple linear regression problems with unequal per unit sampling costs. For. Chron. 62.16-19.

Neter, J.; Wasserman, W. 1974. Applied liner statistical models. Richard D. Irwin, Inc. Homewood, Il. 8 42 p.

Schumacher, F.X.; Hall, F.S. 1933. Logarithmic expression of timber-tree volume. J. Agric. Res. 47: 719-731.

Spurr, S.H. 1952 Forest inventory. Ronald Press, New York, N.Y. 476 p.

Titus, S.J.; Morgan, D.J. 1985. Tree height: Can large-scale photo measurements be more accurate than field measurements? For. Chron. 61: 214-217.

West, P.W.; Ratkowsky, D.A.; Davis, A.W. 1984. Problems of hypothesis testing of regressions with multiple measurements from individual sampling units. For. Ecol. Manage. 7: 207-224.

Verbyla, D. 1986. Potential prediction bias in regression and discriminant analysis. Journ. can. rech. for. 16: 1255-1257.